Hindawi Publishing Corporation Abstract and Applied Analysis Volume 2015, Article ID 823562, 14 pages http://dx.doi.org/10.1155/2015/823562

Research Article

A Data Mining Approach for Visual and Analytical Identification of Neurorehabilitation Ranges in Traumatic Brain Injury Cognitive Rehabilitation

Alejandro García-Rudolph^{1,2,3} and Karina Gibert⁴

Correspondence should be addressed to Alejandro García-Rudolph; alejandropablogarcia@gmail.com

Received 2 January 2015; Accepted 23 February 2015

Academic Editor: Francisco Solis

Copyright © 2015 A. García-Rudolph and K. Gibert. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Traumatic brain injury (TBI) is a critical public health and socioeconomic problem throughout the world. Cognitive rehabilitation (CR) has become the treatment of choice for cognitive impairments after TBI. It consists of hierarchically organized tasks that require repetitive use of impaired cognitive functions. One important focus for CR professionals is the number of repetitions and the type of task performed throughout treatment leading to functional recovery. However, very little research is available that quantifies the amount and type of practice. The Neurorehabilitation Range (NRR) and the Sectorized and Annotated Plane (SAP) have been introduced as a means of identifying formal operational models in order to provide therapists with decision support information for assigning the most appropriate CR plan. In this paper we present a novel methodology based on combining SAP and NRR to solve what we call the Neurorehabilitation Range Maximal Regions (NRRMR) problem and to generate analytical and visual tools enabling the automatic identification of NRR. A new SAP representation is introduced and applied to overcome the drawbacks identified with existing methods. The results obtained show patterns of response to treatment that might lead to reconsideration of some of the current clinical hypotheses.

1. Introduction

Traumatic brain injury (TBI) is a critical public health and socioeconomic problem throughout the world. Although high-quality prevalence data are scarce, it is estimated that in the USA around 5.3 million people are living with a TBI-related disability, and in the European Union approximately 7.7 million people who have experienced a TBI have disabilities [1].

TBI is considered a *silent epidemic*, because society is largely unaware of the magnitude of the problem [2]. The World Health Organization predicts that, by the year 2020, TBI and road traffic accidents will be the third greatest cause of disease and injury worldwide [3].

The consequences of TBI vary from case to case but can include motor, cognitive, and behavioral deficits in the patient, disrupting their daily life activities at personal, social, and professional levels. The most important cognitive deficits after suffering a TBI are those related to attention, decrease in memory and learning capacity, worsening of the capacity to schedule and to solve problems, a reduction in abstract thinking, communication problems, and a lack of awareness of one's own limitations. These cognitive impairments hamper the path to functional independence and a productive lifestyle for the person with TBI.

New techniques of early intervention and the development of intensive TBI care have improved the survival rate noticeably. However, despite these advances, brain injuries

¹Institut Guttmann, Institut Universitari de Neurorehabilitació adscrit a la UAB, Badalona, Barcelona, Spain

²Universitat Autònoma de Barcelona, Bellaterra, Cerdanyola del Vallès, Spain

³Fundació Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Badalona, Barcelona, Spain

⁴Departament d'Estadística i Investigació Operativa, Universitat Politècnica de Catalunya (BarcelonaTech), Jordi Girona 1-3, 08034 Barcelona, Spain

still have no surgical or pharmacological treatment to reestablish lost functions [4]. In this context, cognitive rehabilitation (CR) is defined as a process whereby people with brain injury work together with health service professionals and others to remedy or alleviate cognitive deficits arising from a neurological injury [5].

A typical CR program consists mainly of exercises which require repetitive use of the impaired cognitive system in a sequence of tasks that is progressively more demanding [6]. Both the brain and body need to relearn how to function following neurological injury; harnessing this inherent ability for neuronal circuit change in the brain may be essential if the benefit of rehabilitation is to be maximized.

The process by which neuronal circuits are modified by experience, learning, or injury is referred to as neuroplasticity [7].

While task repetition is not the only important feature, it is becoming clear that neuroplastic change and functional improvement occur after a number of specific tasks are performed but do not occur with other numbers [8, 9]. Thus, one important focus for rehabilitation professionals is the number of repetitions and the type of task performed during treatment. However, there is very little research to quantify the amount and type of practice that occurs during clinical rehabilitation treatment and its relationship to rehabilitation outcomes [10, 11].

In our previous research [12], the *Neurorehabilitation Range (NRR)* was introduced as the conceptual framework by which to describe the degree of performance of a CR task that produces maximum rehabilitation effects. The Sectorized and Annotated Plane (SAP) is proposed as a visual tool to find both the NRR and an operational definition for it, to be used in real clinical practice. Two data-driven methods to build the SAP were introduced in [12] and compared. The NRR of a given task is therefore determined as a rectangular region defined by 2 dimensions: the number of executions of a task during a CR treatment and the performance in each execution of the task.

In this paper, we build on the concept of NRR and SAP tools to solve what we refer to as the Neurorehabilitation Range Maximal Regions (NRRMR) problem. Basically, this consists in the automatic identification of NRRs with datadriven models that are able to avoid the limitations observed in the SAP performance. In the NRRMR, the problem of occlusions that appeared in the SAP, being a pure visualization tool, is overcome, and a variable number of NRRs for a given CR task are found, according to different user-defined conditions concerning the acceptable degree of uncertainty. In the current proposal, the SAP is transformed into a masked binary matrix and a geometric optimization algorithm (the maximal empty rectangle (MER) problem [13]) is generalized to the NRRMR, allowing for the identification of regions satisfying user-defined conditions (see details in Section 3). The proposed methods are extended to any number of tasks grouped in cognitive functions, allowing for the identification of NRR of not only a single task (as in [12]) but also a group of them. Methods are applied in the same real clinical context as in [12] in order to allow comparison of results.

The structure of the paper is as follows: Section 2 briefly presents the state of the art and the starting point of the proposal. Section 3 introduces the proposed analysis methodology and Section 4 its application to the CR context; Section 5 presents a discussion of the obtained results and a comparison with the previous and Section 6 the conclusions and future lines of research.

2. State of the Art

There is a common belief that CR is effective for TBI patients, based on a large number of studies and extensive clinical experience. Different statistical methodologies and predictive data mining methods have been applied to predict clinical outcomes of the rehabilitation of patients with TBI [14–16]. Most of these studies focus on determining survival, predicting disability or the recovery of patients, and looking for the factors that better predict the patient's condition after TBI.

However, current knowledge about the factors that determine a favorable outcome is mainly empirical and the benefit of such interventions is still controversial [17] (see also Ecri Cognitive Rehabilitation Therapy for Traumatic Brain Injury: What We Know and Don't Know about Its Efficacy. Editorial Note 10/11/11: IOM's New Report on Brain Injury Treatments Draws Conclusions Similar to ECRI Institute's Earlier Findings). The development of new tools to evaluate scientific evidence of such effectiveness will contribute to a better understanding of CR.

It seems that patient improvement might depend *inter alia* on the location of the injuries, cognitive profile, duration, and intensity of proposed treatments and their level of completion [18, 19]. However, these seem to be only some of the determining factors and they cannot by themselves explain the overall phenomenon. Although these factors are considered in the design of rehabilitation treatments, other relevant factors exist that are much more difficult to control and which are related to the high variability of the lesions, the complexity of cognitive functions, and the lack of proper instrumentation by which to systematize interventions. This produces intrinsic group heterogeneity and the classical comparative studies do not perform well [20], which makes it difficult to advance the pathophysiology of cognitive neurorehabilitation knowledge.

In [21], basic machine learning, algorithms were used to predict the probability of improvement in a patient according to their initial neuropsychological assessment. This approach was able to identify subpopulations of patients more suitable for improvement using CR treatments. However, it did not provide any information to help CR therapists adapt CR programs to increase the improvement itself or to enlarge the subpopulations that might activate improvement to CR treatments. Going a little bit further, in [22] the performance obtained by the patient in a certain task has been included in the model together with the initial assessment. Machine learning methods significantly improved predictive capacity. This work provided evidence that task performance is involved in patient improvement. However, it did not provide

information on successful patterns of tasks to be proposed to the patients leading to improvements in cognitive functions.

For these reasons, other approaches have to be found to better understand the CR process, with the aim of obtaining scientific evidence about its effectiveness and providing relevant information for the establishment of general guidelines for CR program design that can assist CR therapists in clinical practice. Analyzing data from new perspectives can contribute to this field [23].

Our proposal in [12] approaches the problem from a datadriven perspective by developing new data mining tools that can reduce uncertainty in the field. The paper introduces elements to assess when a patient is performing a task under a Neurorehabilitation Range, as an indicator that maximum improvement of the patient might be expected on the targeted cognitive function. This contributes to a better understanding of the role over clinical improvement of a particular degree of performance of a CR task. Two different methodologies to build the SAP were proposed in [12]: direct construction by visualization of raw data (Vis-SAP method) and DT-SAP, which is based on decision tree induction and therefore could be automated. Decision trees have been considered because their inherent structure leads directly to the NRR model. This is built as the OR of all branches leading to a leaf labeled as improvement. Both methods effectively determine the areas where the probability of improvement is higher; a statistical two-proportion test has been used to assess the quality of the NRR models by checking whether the probability of improvement is significantly higher when tasks are performed according to NRR than away from it. Whereas DT-SAP is a deterministic method that can be automated, the Vis-SAP is a semideterministic method that requires visual inspection as a final step. However, it seems to produce better results in practical applications because the incomplete sectorization of the plane in very homogeneous areas provided by Vis-SAP outperforms the results induced from a DT where the leaves are often contaminated; that is, they contain both improving and nonimproving patients. However, the Vis-SAP method has a limitation: the graphical representation used in [12] does not take into account the occlusions. This means that whether or not a pixel in the graph is depicted as an improvement depends on the majority of pixel points but this does not take into account the error tax produced by this simplification. In this paper, a new methodology is provided that overcomes this limitation.

2.1. Maximal Empty Rectangle. The key idea of the present work is to transform the Vis-SAP method from [12] into a geometric optimization algorithm that avoids the visual effect of occlusions while permitting some degree of impurity in the detected areas of the NRR to be taken into account.

For this purpose, a generalization of the MER problem will be introduced. The MER problem consists of recognizing all maximal empty axes-parallel (isothetic) rectangles, in a rectangular space region where some points are located. It was first introduced in 1984 [13] as follows.

Given a rectilinearly oriented rectangle *A* in the Cartesian plane and *a* set $S = \{P_1, P_2, ..., P_n\}$ of n > 1

points in the interior of A, where each point P_i is specified by its X and Y coordinates (X_i, Y_i) , $i = 1, 2, \ldots, n$, and A specified by its left boundary A_l , right boundary A_r , top boundary A_t , and bottom boundary A_b , the maximum empty rectangle (MER) problem is to find a maximum area rectangle whose sides are parallel with those of A and which is contained in A such that no point of S lies in its interior.

Several algorithms have been proposed for the planar problem over the years [24]. For instance, an early algorithm by Chazelle et al. [25] runs in $O(n\log^3 n)$ time and $O(n\log n)$ space. The fastest known algorithm, proposed by Aggarwal and Suri in 1987 [26], runs in $O(n\log^2 n)$ time and O(n) space. A lower bound of $\Omega(n\log n)$ in the algebraic decision tree model for this problem has been shown by McKenna et al. [27].

This problem arises in situations where a rectangular shaped plant is to be located within a similar region which has a number of forbidden areas or when a "perfect" rectangular piece from a large similarly shaped metal sheet with some defective spots [13] is to be cut. The problem could also be further modified so that the length and width of the sought-after rectangle have a certain ratio or a certain minimum length.

Maximal empty rectangles also arose in the enumeration of *maximal white rectangles* in image segmentation [28].

More recently, applications can be found in data mining [29], geographical information systems (GIS), and very large-scale integration design [30].

To the best of our knowledge, MERs have not yet been applied, either for NRR identification in particular or in the context of CR in general.

3. Materials and Methods

The proposed methods present two strategies for the analytical and graphical identification and visualization of NRR and non-NRR based on the notion of SAP as introduced in [12] and on the classical MER problem, respectively.

3.1. Sectorized and Annotated Plane (SAP). Given three variables Y, X1, and X2, where Y is a qualitative response variable, with values $\{y1, y2, \ldots\}$, and X1, X2 numerical explanatory variables, the SAP is a 2-dimensional plot with X1 in the x-axis, X2 in the y-axis and rectangular regions with constant Y displayed and labeled with Y values as outlined in Figure 1. An SAP is therefore a graphical support tool aimed at visualization, where the response variable is constant in certain regions of the $X1 \times X2$ space. Eventually, allowing a relaxation of strict constant Y in the marked regions, the SAP might include an indicator of region purity, adding the probability of occurrence of the labeling value.

Given a particular CR task and assuming *Y* as a binary variable reporting improvement of the patient in the cognitive function targeted by the task (*YES*, *NO*), the SAP leads to response zones where participants show similar response

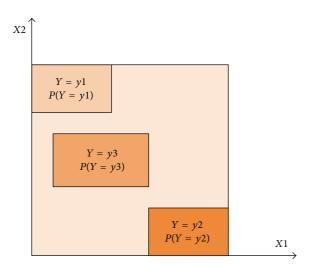


FIGURE 1: General Sectorized and Annotated Plane (SAP) description.

to treatment. The SAP shows a plane sectorization directly related to treatment response. This allows identification of logical restrictions (rules) determining different treatment outcomes.

3.2. Visualization-Based SAP (Vis-SAP). Data is plotted regarding X1 and X2, and each point is marked with different colors according to the values of Y. This categorized scatterplot (sometimes known as a letterplot) is an exploratory technique for investigating relationships between X1 and X2 within the subgroups determined by Y. For the particular application presented here, X1 is the result obtained at every single execution (e.g., an integer number in the range $0\cdots 100$) and X2 is the number of executions of the task performed by the subject, while Y is the effect of the neurorehabilitation process (improvement/nonimprovement).

This exploratory analysis is used to identify systematic relationships between variables when there is no previous knowledge about the nature of those relationships. The constant *Y* regions detected in the plot can be expressed in the form of logical rules involving the implied variables in the following form:

if (Result in [r1, r2] and Executions in [R1, R2]) then P (Improvement) = p,

where r1, r2, R1, and R2 indicate the limits of the regions detected in the graph.

The SAP is built on the basis of these rules.

3.3. Frequency Table SAP (FT-SAP). The main problem with Vis-SAP is that in every pixel in the image several points might be overlapped and not always labeled with the same response value. Detection of NRR regions is performed by labeling each pixel with the majority label, using a simple voting scheme and without taking into account the balance between improvement and nonimprovement pixels overlapped.

Table 1: p_{ij} is the proportion of improving patients in pixel (i, j).

Execs./Results	0	1	• • •	i
1				
2				
:				
j				$p_{ij} = m_{ij}/n_{ij}$
•••				

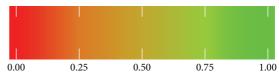


Figure 2: Color gradient for p_{ij} values in quartiles.

In this work the first idea proposed is to use a numerical representation of the Vis-SAP based on a two-way matrix, precisely indicating how many points of each class are overlapped at any pixel in the graph.

As in Vis-SAP, in this approach *X*1 is the result obtained at every single execution (Result), *X*2 is the number of executions of the task performed by the subject (Executions), while *Y* is the effect of the neurorehabilitation process: improvement/nonimprovement (*YES*, *NO*) assessed by standardized neuropsychological tests.

Given *mExec* the maximum number of Executions of a task and *mResults* the maximum scoring of a task i = (1 : mExec), j = (1 : mResults), we define

 m_{ij} = number of subjects such that (X2 = i) & (X1 = j) & (Y = YES),

 n_{ij} = number of subjects such that (X2 = i) & (X1 = j),

 $p_{ij} = m_{ij}/n_{ij}$ = percentage of subjects such that (X2 = i) & (X1 = j) & (Y = YES).

For each (i, j) the matrix $P = (p_{ij})$ is built (see Table 1).

An FT-SAP is a graphical visualization where a gradient color from red to green can be assigned to pixel (i, j) according to its p_{ij} as shown in Figure 2.

Given a threshold $\gamma \in [0,1]$ the NRRMR regions can be found over the FT-SAP (γ) as the set of regions $(r,s) \times (t,u)$ such that

$$\forall (i, j) \quad r \leq i \leq s, \quad t \leq j \leq u, \quad \text{it holds } p_{ij} \geq \gamma.$$

Given γ , a 2-color gradient can be defined providing a neat heatmap of the FT-SAP (γ) (as shown in Figure 3). q_{ij} is defined as

$$q_{ij} = \begin{cases} 1, & p_{ij} \ge \gamma, \\ 0, & p_{ij} < \gamma. \end{cases}$$
 (2)

Therefore a binary matrix Q is obtained, with $Q = (q_{ij})$. Q is a mask over FT-SAP filtering pixels according to γ (for empty cells, no color is provided for the pixel).

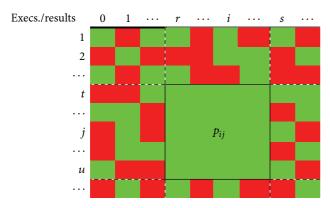


FIGURE 3: A 2-color heatmap defining a two-dimensional NRR for $p_{ij} \ge \gamma$.

3.4. Analytical Identification of NRR. Taking as input parameter the Q matrix resulting from filtering FT-SAP over γ , a method to automatically identify NRR (given maximum width and length of the surface to be searched as user-defined parameters) is described below. The idea is to find all rectangular groups of 1s cells equal to or greater than the minimum width and length provided by the user.

It is solved by two-pass linear O(n) time algorithm (n being the number of cells in the input matrix). As shown in Figure 4, first pass scans the matrix by columns, numbering cells consecutively until a red element (a 0 cell) is found, and second pass scans by rows, searching for elements matching the length and width provided as parameters.

As is shown in the R code in Algorithm 2 the method allows for the simultaneous identification of the NRRs satisfying the user-defined conditions. The MAXRES and MAXEXEC values in the R code correspond to *mResults* and *mExec*, respectively, as defined above. The proposed pseudo code is introduced in Algorithm 1.

With Algorithm 2, the green rectangles as specified by the user in the FT-SAP for a given threshold γ can be identified and NRR established accordingly.

As will be seen in the application section, some real cases provide large green areas contaminated by a small percentage of isolated red points that could be assumed as part of the NRR, provided that an uncertainty tax becomes associated with it. This implies modification of the previous algorithm to find regions with a certain degree of contamination. But the generalization about the provided implementation is not evident. Thus, a classical version of the MER algorithm has been used instead and properly modified. Section 3.5 provides our implementation (Algorithm 3) of the classical MER and Section 3.5.1 provides the proposed generalization to permit a certain degree of contamination in the regions (Algorithm 4).

3.5. Maximal Empty Rectangle (MER) Method. As a first attempt the direct approach to the MER problem is followed: Scan through the matrix, stopping at each element. Treat each element as a potential top-left corner of the MER. For each such top-left corner, try all other elements as

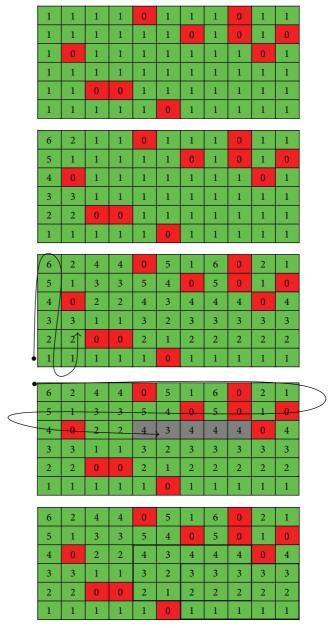


FIGURE 4: Example of the two-pass algorithm.

a potential bottom-right corner of the MER. This approach is implemented in Algorithm 3.

Regarding the performance, in this approach each top-left corner visits about O(mn) locations. For each such top-left corner, the bottom-right corner visits no more than O(mn) positions. An evaluation (checking for 1s) takes O(mn) in the worst-case for each rectangle checked (total: $O(m^3n^3)$) (worst-case). This means that finding pure regions in the Q matrix performs better over time when Algorithm 2 is used. Some improvements to the classical MER approach have been identified which improve performance: checking the area first before scanning for 1s and prune, ignoring the rectangle when the area is too small, also, eliminating as many size 1

```
Input
Anxm matrix of red(0)/green(1) elements obtained after FT-SAP (\gamma):
MAXROW maximum number of rows
MAXCOL maximum number of columns
Output
NRR maxrowxmaxcolumn
First pass
For each column from bottom to top
  Repeat
   Number green element incrementally
  Until a red element is found → Restart numbering
Second pass
For each row from left to right
 NRRrows = 0 #Number of rows of the NRR solution so far
 Repeat
 If element >= MAXCOL
     Increment NRRrows
     NRR = NRR + NRR[element]
 Else NRRrows = 0
 Until NRRrows = MAXROW
Return NRR
```

Algorithm 1

```
#First pass
for (j \text{ in } 1 : MAXRES){
cont \leftarrow 1
for (i \text{ in MAXEXEC} : 1) {
if (y[i, j] == 0){
cont \leftarrow 1
b[i, j] \leftarrow NA
else \{b[i, j] \leftarrow \text{cont}
cont \leftarrow cont + 1
#Second pass
Mdat \leftarrow b
apply(mdat, 1, function(x) {
 r \leftarrow \text{rle}(x \ge \text{MAXROW})
 w \leftarrow \text{which}(\text{!is.na}(r\text{values}) \& r\text{values} \& r\text{lengths} >= \text{MAXCOL})
  if (length(w) > 0) {
    lapply(w, FUN = function(w1){before \leftarrow sum(r$lengths[1:w1]) -r$lengths[w1];
                           c(before + 1, before + r lengths[w1]))
 } else
    NULL
})
```

Algorithm 2

rectangles from the search as possible and checking corners for 0 s before proceeding.

3.5.1. Neurorehabilitation Range Maximal Regions (NRRMR) Problem. To allow for the identification of nonempty regions (i.e., regions containing some degree of 0 values) a modification of the *checkFilled* function is introduced as shown in Algorithm 4. A user-defined tolerance is included as input

to the function and only when that value is exceeded is the area considered as not filled. Figure 5 shows the identification of the maximal rectangle containing one nonempty element as output ([topLeftx, topLefty, botRightX, botRightY] = [5, 1, 7, 8] area = 24), instead of the bottom-right rectangle that would be the output if no tolerance parameter is introduced ([topLeftx, topLefty, botRightX, botRightY] = [13, 8, 16, 12] area = 20).

```
Amxn matrix of red/green elements obtained after FT-SAP(\gamma)
Output
MER submatrix of A
findMaxRectangleArea \leftarrow function (A) \{
# (1) Initialize.
maxArea \leftarrow 0;
area \leftarrow 0:
# (2) Outer double-for-loop to consider all possible positions for top-left corner.
for (i \text{ in } 1 : m){
    for (j \text{ in } 1 : n){
   # (2.1) With (i, j) as top-left, consider all possible bottom-right corners.
    for (a \text{ in } i : m){
   for (b \text{ in } i : n){
    # (2.1.2) See if rectangle(i, j, a, b) is filled.
            Filled \leftarrow checkFilled (i, j, a, b);
  # (2.1.3) If so, compute it's area.
              if (filled){area \leftarrow computeArea (i, j, a, b)}
        # If the area is largest, adjust maximum and update coordinates.
                    if (area > maxArea){
                                 maxArea ← area;
                                 topLeftx \leftarrow i;
                                \mathsf{topLeft} y \leftarrow j;
                                 botRightX \leftarrow a;
                                 botRightY \leftarrow b
\max R \leftarrow c(\text{topLeft}x, \text{topLeft}y, \text{botRight}X, \text{botRight}Y)
return (list(area = \max Area, rect = \max R));
computeArea \leftarrow function (i, j, a, b) {
if (a < i) {return(-1)}
if (b < j) {return(-1)}
return ((a - i + 1) * (b - j + 1))
checkFilled \leftarrow function (i, j, a, b) {
for (k1 \text{ in } i : a){
    for (k2 \text{ in } j:b){
        if (A[k1, k2] == 0) {return (FALSE)}
return (TRUE)
```

Algorithm 3

4. Application and Results

4.1. Clinical Context. This work is based on the same context as in [12], the Neuropsychological Department of Institut Guttmann Neurorehabilitation Hospital (IG). The Information Technology framework for CR treatments in this clinical setting is therefore the PREVIRNEC© platform [31]. It is specifically designed to operate CR plans assigned to subjects, as well as to manage precise follow-up information about the process.

Three main cognitive functions are usually addressed in a CR program [6]: attention, memory, and executive functions;

all of them can profoundly affect individuals' daily functioning. Even mild changes in the ability to attend, process, recall, and act upon information can have a significant effect on the quality of life of the patient.

Before starting the CR program at IG every patient undergoes a Neuropsychological Assessment Battery (NAB). This battery includes 28 items covering the major cognitive domains (attention, memory, and executive functions) measured using standardized cognitive tests.

Differences between pre- and posttreatment NAB test scores are used to measure particular patient improvement

```
checkFilled ← function (i, j, a, b, TOLERANCE) {
  tol ← 0;
  for (k1 in i : a) {
    for (k2 in j : b) {
        if (A[k1, k2] == 0) {
        tol ← tol + 1;
        if (tol > TOLERANCE) {return (FALSE)}}
    }
}
return (TRUE)
}
```

Algorithm 4

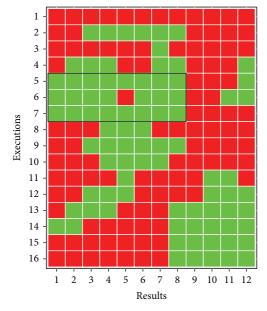


FIGURE 5: MER with user-defined tolerance = 1.

in the fields of attention, memory, and executive functions. Improvement criteria in the respective cognitive functions are defined in IG cognitive rehabilitation protocols.

For each patient the therapist creates a specific CR treatment, that is, a sequence of tasks. At IG a typical CR program in PREVIRNEC© platform ranges from 2 to 4 sessions a week for a period of 2 to 5 months. After the execution of a given task the patient gets a result ranging from 0 to 100: a 0 result denotes the lowest level of task completion and a 100 the highest. At the moment of this analysis PREVIRNEC© platform supports 96 different CR tasks targeting the three main cognitive functions mentioned above (17 regarding attention rehabilitation, 59 memory, and 20 executive functions). In a typical CR treatment every patient executes a different number of tasks in a different order; the same task could be executed *n* times by a patient and may not be included during the whole treatment of another patient, depending on the decision of the therapist.

One hundred and twenty-three TBI adults following a 3- to 5-month CR treatment at IG Neuropsychological

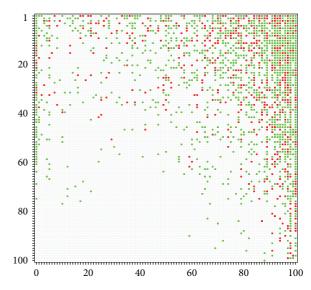


FIGURE 6: FT-SAP(0.8) for idTask = 146 for a total number of 3329 executions.

Rehabilitation Unit are analyzed in this study. For every patient the following demographic and clinical variables are considered: age, gender, educational level, Glasgow Comma Scale (GCS), and posttraumatic amnesia (PTA) duration. Table 2 shows the basic statistics for numerical variables.

Initial assessment of the TBI severity is reported according to GCS levels. A GCS score of eight or less after resuscitation from the initial injury is classified as a *severe* brain injury. The GCS score for *a moderate* brain injury ranges between nine and thirteen and a score of thirteen or higher indicates a *mild* brain injury or concussion. For the patients analyzed, most GCS scores (86.17%) show *severe* brain injury level (mean value 6.45 ± 3.15).

The following methods have been implemented and executed in R version 2.15.1 (2012-06-22), "Roasted Marshmallows" Copyright © 2012 (the R Foundation for Statistical Computing, ISBN 3-900051-07-0, Execution Platform: x86_64-pc-mingw32/x64 (64-bit)).

4.2. Visual Identification of NRR Considering One Task. The first application is the FT-SAP presented in Section 3.3 for a CR task (idTask = 146 targeting the attention cognitive function) with γ = 0.8. The 2-color heatmap shown in Figure 6 is obtained. "Results" are plotted along the x-axis ranging from 0 to 100 and "number of executions" along the y-axis, also ranging from 0 to 100. Two neat NRR regions can be visually identified for high values of Result and mid to high values of number of executions. The identified NRR might indicate that other tasks of the same type (e.g., targeting the same function or subfunction) could behave in a similar way; Section 4.4 below shows results for tasks grouped by cognitive function.

CR treatment for this task comprises 3329 executions in total, where 1950 of them correspond to patients with improvement = YES and 1379 to improvement = NO.

Variable	N	N^*	Mean	Std. Dev.	Min	Q1	Median	Q3	Max
AGE	123	0	36.56	6.50	18	25	32	40	68
GCS	89	34	6.45	3.15	0	4	6.5	40	14
PTA	40	83	131.6	140.5	34	79	103	136	947

TABLE 2: Basic descriptive statistics for numerical variables.

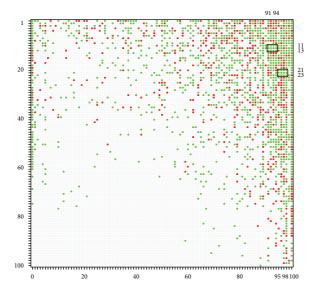


FIGURE 7: NRR coordinates identified by the proposed method.

4.3. Analytical Identification of NRR. The method presented in Section 3.4 is applied for the analytical identification of the NRRs. The results obtained are shown in Figure 7 with input parameter values MAXROW = 4 and MAXCOLUM = 3.

The resulting NRRs are the following:

if (Results in [91, 94] and Repetitions in [11, 13]) then P (Improvement) ≥ 0.8 ,

if (Results in [95, 98] and Repetitions in [21, 23]) then P (Improvement) ≥ 0.8 .

4.4. Visual Identification of NRR Considering Every Task by CR Function. As introduced in Section 4.1, this study analyzes one hundred and twenty-three TBI adults following a 3-5month CR treatment at the IG Neuropsychological Rehabilitation Unit. PREVIRNEC© platform includes 17 tasks addressing the attention function, 59 memory, and 20 executive functions. During this CR treatment, the total number of task executions is 41010 (15475 targeting attention, 14557 memory, and 10978 executive functions). Figure 8 shows FT-SAP ($\gamma = 0.8$ left column and $\gamma = 0.9$ right column) for every execution of tasks grouped by CR functions. The top pair of plots corresponds to the execution of attention tasks, the middle pair to memory tasks, and the bottom pair to executive functions. Three different responses to CR treatment patterns can be identified according to how improvement points are distributed. Attention tasks are grouped on medium to high values of Results and medium to low values of number of executions. Memory is more uniformly spread from low to

high values of results; executions are all over the plot and executive functions are a mix of the above patterns with concentration on high values and also for specific lower values of results and executions.

4.5. Analytical Identification of NRR (MER Method). The methods presented in Sections 3.5 and 3.5.1 are applied for the analytical identification of NRRs.

The first plot in Figure 8 (attention tasks with $\gamma=0.8$) is now analyzed using the method presented in Section 3.5.1 to identify maximum zones of improvement for every execution of attention tasks, allowing for a tolerance of 2 elements. Obtained results (graphically represented in Figure 9) are as follows:

 $\gamma = 0.8$, [topLeftx, topLefty, botRightX, botRightY] = [11, 87, 20, 88], area = 20, tolerance = 2: $\gamma = 0.8$, [topLeftx, topLefty, botRightX, botRightY] = [16, 98, 25, 100], area = 30,

leading to the following NRRs:

if (Results in [87, 88] and Repetitions in [11, 20]) then P (Improvement) ≥ 0.8 ,

if (Results in [98, 100] and Repetitions in [16, 25]) then P (Improvement) ≥ 0.8 .

5. Discussion

This work aims to identify the conditions in which performing a certain cognitive rehabilitation task (or a group of tasks) guarantees better potential for the activation of brain plasticity and therefore helps bring about improvements in the assessed cognitive functions after CR treatment. As this research takes our previous research as a starting point, the results comparison is provided below and the pros and cons are discussed.

Figure 10(a) presents FT-SAP proposed in this work for idTask = 151 and γ = 1 and Figure 10(b) shows Vis-SAP obtained in [12]. As presented in Section 3.3, FT-SAP(1) represents a green point at position (i, j) if $p_{ij} \ge \gamma$, where p_{ij} = 1; that is, all patients executing i times Task 151 and obtaining score j improve after treatment. In Figure 10(a), gray cells do not register observations. As shown in Figure 10(b), no

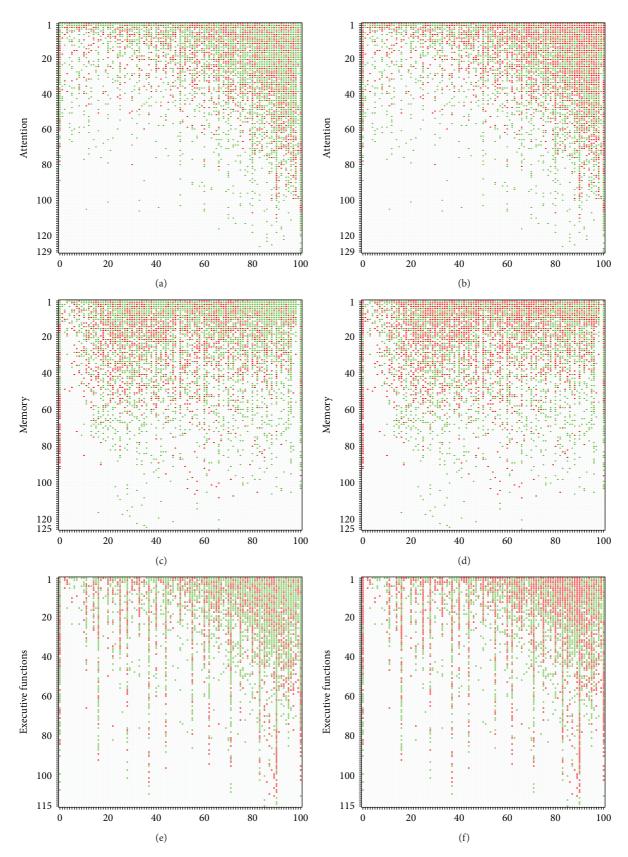


Figure 8: FT-SAP for each cognitive function: attention (top), memory (middle), and executive functions (bottom), γ = 0.8 left column and γ = 0.9 right column.

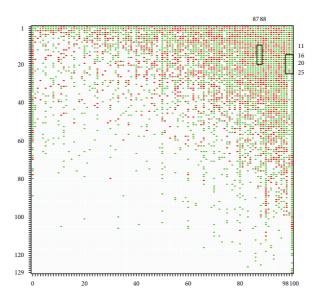


FIGURE 9: Analytical identification of NRR with and without user-defined tolerance.

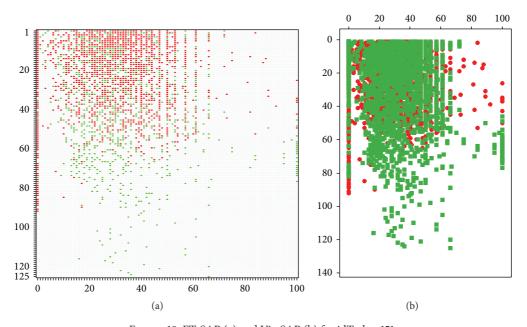


Figure 10: FT-SAP (a) and Vis-SAP (b) for idTask = 151.

subject with Y = NO executed the task more than 60 times obtaining results other than zero, leading to the identified rule: (NRR(151) = Execs151 > 65 and Res > 20).

However, it can be seen that the area $[22,35] \times [15,30]$ appears as a totally green area in the Vis-SAP, whereas there are plenty of red points in the FT-SAP. This indicates that most of the points in this area do not have 100% of patients improving. This is the major contribution of the FT-SAP. One can evaluate the degree of certainty of the induced NRR as the points occlusion occurring in the VIS-SAP is overcome. On the other hand, in the areas of the plot with high concentrations of executions and results (as shown for results lower than 40 and number of executions lower than 60 in Figure 10(b)) Vis-SAP does not provide a neat visualization.

By construction, FT-SAP avoids the confusion produced by overlapping points. Decreasing γ to 0.5, that is, admitting half of the patients in a nonimproving point after the treatment, produces an FT-SAP(0.5) as shown in Figure 11 with many more green points, but it is still difficult to identify an interesting rectangular green region to establish a second area of NRR for Task 151. In conclusion, the FT-SAP provides a refinement of the Vis-SAP that enables uncertainty to be dealt with and avoids, by construction, confusions produced by several patients overlapping in the same point.

When longer periods of CR treatments are considered, including therefore an increasing number of subjects, the areas of the plot where no task executions can be found tend to decrease. Also, when a group of tasks targeting the same

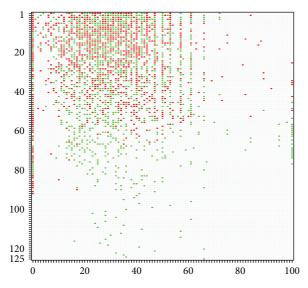


FIGURE 11: FT-SAP(0.5) for idTask = 151.

cognitive function is considered instead of a single one, more robust NRR can be induced from the proposed FT-SAP representation.

In addition, both plots in Figure 10 agree on the identification of a zone where NRR is not achieved, as shown by the high values for results and the low number of executions. This seems to suggest that, for this type of task, the therapist might expect to achieve NRR for lower results. An explanation for this might be that the rehabilitating effect of a task depends on the ratio between the skills of the treated patient and the challenges involved in the execution of the task itself. The difficulty is related to the level of stimulation of cognitively involved functions; maximum activation occurs when the task is "just barely too difficult" [32]. If the task is either too easy or too difficult for the patient, it appears to be less effective. Active monitoring of the subject's progress is therefore required to adapt the difficulty of the tasks to the potential capacities and progress of the subject, always pushing them to reach a goal just beyond what they can attain, but not too far. Thus, determining the correct training schedule requires a very precise tradeoff between sufficiently stimulating and sufficiently achievable tasks, which is far from intuitive, and is still an open problem, both empirically and theoretically.

At the moment of submission, PREVIRNEC® was assuming as clinical hypothesis a constant NRR for the whole set of available tasks. The assumed NRR is a one-dimensional NRR which only takes into account the scorings obtained in the execution of tasks. Thus, a task is considered to be executed in NRR if the scoring obtained falls in the interval [65, 85]. Therefore the PREVIRNEC® system automatically increases the difficulty if the patient performs beyond the NRR (i.e., achieving a result higher than 85, meaning that the task was too easy for the patient, and thus stimulating the required brain areas only to a poor degree) and decreases it if he/she performs below NRR (meaning that the effort demanded for the task was so hard that it became impossible,

and thus nontherapeutic effects can be achieved). The current proposal enables a more precise refinement of the system where the NRR might change from task to task, depending on its own characteristics and the type of stimulation involved.

As shown in Section 4.4 where tasks are analyzed grouped by cognitive function, a different pattern can be identified for groups of tasks. Executive functions (at the bottom of Figure 8) seem to be a combination of the attention and memory plots. An explanation for this might be that executive functions are those abilities that allow individuals to efficiently and effectively engage in complex goal-directed behaviors such as planning, sequencing, categorization, flexibility, and inhibition. According to Lezak [33], this includes the capacity to set goals, to form plans, to initiate actions, and to regulate and evaluate behavior according to the plan and to situational constraints. Therefore executive functions are considered higher level functions which control the more basic cognitive functions such as attention and memory. This implies that intactness of the executive functions might determine whether a brain-damaged individual with lower level cognitive deficits, for example, selective or divided attention processing or memory deficits, is able to compensate for these deficits and to adapt to the altered situations by restructuring activities [34].

This suggests that the current NRR considered (the scoring interval [65, 85]) might be enlarged to also include the number of executions as introduced in [12] and also could be addressed by cognitive function, possibly leading to a different NRR for each function, as shown in Section 4.4.

The methods presented in Section 3 were tested on a Windows 7 Professional SP 1 PC, Intel Core i3 2.40 GHz (2 GB RAM) 64-bit OS.

The algorithm presented in Section 3.4 runs in few seconds. The MER method described in Sections 3.5 and 3.5.1 took about 15 minutes to execute. Though inefficient, it provides a good basis upon which to build. To improve its performance, direction to the search needs to be introduced. The proposed algorithm could enumerate the subrectangles in any random order and still find the correct solution. Instead, we might take advantage of the fact that if a small rectangle contains a zero, so will each of its surrounding rectangles. Therefore, rectangles will be grown for each possible lower-left corner. This growing process will only produce upper-right corners defining rectangles which contain only successes (ones, i.e., improvements).

As presented in [12] the main drawbacks of the Vis-SAP proposal are twofold: on the one hand, the lack of completeness of the Vis-SAP criterion proposed. Indeed, looking at the SAP diagram, VIS-SAP is not assigning improvement or nonimprovement to the whole area, but only to small parts of the diagram corresponding to concrete and reduced areas where either improvement or nonresponse can be ensured. Therefore it could be said that Vis-SAP provides a semideterministic procedure where a particular configuration for both results and repetitions ensures improvement, a second configuration where the task does not produce patient improvement, and out of these regions the outcome is undetermined. On the other hand, the proposed analysis considers each task individually, being NRR defined for every

single task. FT-SAP and the proposed NRRMR methods overcome both these drawbacks.

6. Conclusions and Future Work

This work builds on our previous contribution towards the design, implementation, and execution of personalized, predictable, and data-driven CR programs. We wish to identify NRR for cognitive rehabilitation tasks that lead to patient improvement.

SAP and MER problems were used to automatically generate data-driven models in order to identify bidimensional NRRs, taking into account the proper combinations of repetition of tasks and performance. In this work, a new SAP is proposed overcoming the identified limitations of Vis-SAP and allowing for the automatic identification of a bidimensional NRR for a given task. A method is introduced to identify a variable number of NRRs satisfying a certain degree of reliability (γ) for a given task. A direct MER algorithm is implemented and modified to identify regions of minimum probability of improvement γ , in order to solve the Neurorehabilitation Range Maximal Regions (NRRMR) problem introduced in this paper. The proposed methods are also applied to any number of CR tasks grouped in cognitive functions allowing for the identification of NRR, not only for a single task but also for a group of them stimulating the same cognitive function. When grouped by cognitive functions, a different response pattern has been identified for memory skills, attention, or executive functions, suggesting that NRR might also depend on the targeted function. Further analyses, including subfunctions of each cognitive function, are currently underway. In PREVIRNEC© platform each CR task is designed to target a cognitive subfunction (e.g., idTask 151 analyzed above targets the visual memory subfunction of the memory function) and the improvement/nonimprovement values of Y variable can therefore be determined by the specific subtests of the NAB assessment which evaluate that subfunction, leading to a finer granularity of the results.

Analytical and visual tools are proposed, designed, implemented, and executed to find an operational approach for the identification of a bidimensional NRR from a data-driven approach. FT-SAP has been introduced as a parametric heatmap-based visualization tool to find areas where a target event has a minimum probability of occurring. For this particular application, the FT-SAP identifies areas with high probability of cognitive improvement. Although FT-SAP is not a complex concept, it has shown great potential for finding the NRR region of a cognitive rehabilitation task (or a set of tasks) in an automatic, efficient, simple, and very intuitive way. Identified NRRs will be validated with a random group of patients not included in this analysis to verify the obtained results.

As a complementary visual and analytical tool, starting from the representation provided by FT-SAP, the MER problem method has been introduced in order to identify maximum NRR. An existing MER method has been implemented and adapted to support a user-defined tolerance to the search for MER. This allows the therapist to define

the extent to which the MER should be empty. The tolerance specifies the number of nonempty elements allowed in an MER. The visual identification of these regions to allow a certain degree of nonempty elements is not straightforward and therefore an automatic identification provides therapist with this additional information. As suggested in Section 3.5, the current NRRMR implementation is due to be improved with regard to faster computational time. The tolerance parameter can also be adapted to be a percentage of points of the identified area instead of a fixed number of points, thus supporting therapists with more elements for a potentially good response to treatment.

As presented in Section 1, other factors are supposed to be highly determinant of response to treatment, such as the TBI severity reported by GCS, the time since injury, age, and educational level [20]. Extension of the current proposals to include such other factors is currently being explored, provided that the formal framework of FT-SAP is easily extendable to hypercubes instead of two-way tables as shown in this work.

Conflict of Interests

No competing financial interests exist.

Acknowledgments

This research was supported by the Ministry of Science and Innovation (Spain) INNPACTO Program (PT NEUROCONTENT, Grant no. 300000-2010-30), Ministry of Education Social Policy and Social Services (Spain) IMSERSO Program (PT COGNIDAC, Grant no. 41/2008), MARATÓ TV3 Foundation (PT: Improving Social Cognition and Meta-Cognition in Schizophrenia: A Tele-Rehabilitation Project, Grant no. 091330), EU CIP-ICT-PSP-2007-1 (PT: CLEAR, Grant no. 224985), Spanish Ministry of Economy and Finance (PT COGNITIO, Grant no. TIN2012 38450), and EU-FP7-ICT (PT PERSSILAA Grant no. 610359).

References

- [1] B. Roozenbeek, A. I. R. Maas, and D. K. Menon, "Changing patterns in the epidemiology of traumatic brain injury," *Nature Reviews Neurology*, vol. 9, no. 4, pp. 231–236, 2013.
- [2] J. A. Langlois and R. W. Sattin, "Traumatic brain injury in the United States: research and programs of the Centers for Disease Control and Prevention (CDC)," *Journal of Head Trauma Rehabilitation*, vol. 20, no. 3, pp. 187–188, 2005.
- [3] S. A. Tabish and N. Syed, "Traumatic brain injury: the neglected epidemic of modern society," *International Journal of Science and Research*, vol. 3, no. 12, 2014.
- [4] D. T. Stuss, G. Winocur, and I. H. Robertson, Cognitive Neuro-Rehabilitation: Evidence and Application, Cambridge University Press, Cambridge, UK, 2nd edition, 2008.
- [5] B. A. Wilson, "La réadaption cognitive chez les cérébro-lésés," in Neuropsychologie Clinique et Neurologie du Comportement, M. I. Botez, Ed., pp. 637–652, Les Presses de l'Université de Montreal, Montreal, Canada, 2nd edition, 1996.

- [6] M. M. Sohlberg, Cognitive Rehabilitation. An interactive Neuropsychological Approach, 2001, edited by: C. A. Mateer.
- [7] R. J. Nudo, "Adaptive plasticity in motor cortex: implications for rehabilitation after brain injury," *Journal of Rehabilitation Medicine*, no. 41, supplement, pp. 7–10, 2003.
- [8] J. R. Carey, W. K. Durfee, E. Bhatt et al., "Comparison of finger tracking versus simple movement training via telerehabilitation to alter hand function and cortical reorganization after stroke," *Neurorehabilitation and Neural Repair*, vol. 21, no. 3, pp. 216– 232, 2007.
- [9] S. L. Wolf, C. J. Winstein, J. P. Miller et al., "Effect of constraint-induced movement therapy on upper extremity function 3 to 9 months after stroke: the EXCITE randomized clinical trial," *Journal of the American Medical Association*, vol. 296, no. 17, pp. 2095–2104, 2006.
- [10] C. K. English, S. L. Hillier, K. R. Stiller, and A. Warden-Flood, "Circuit class therapy versus individual physiotherapy sessions during inpatient stroke rehabilitation: a controlled trial," *Archives of Physical Medicine and Rehabilitation*, vol. 88, no. 8, pp. 955–963, 2007.
- [11] S. Kuys, S. Brauer, and L. Ada, "Routine physiotherapy does not induce a cardiorespiratory training effect post-stroke, regardless of walking ability," *Physiotherapy Research International*, vol. 11, no. 4, pp. 219–227, 2006.
- [12] A. García-Rudolph and K. Gibert, "A data mining approach to identify cognitive NeuroRehabilitation Range in Traumatic Brain Injury patients," *Expert Systems with Applications*, vol. 41, no. 11, pp. 5238–5251, 2014.
- [13] A. Naamad, D. T. Lee, and W.-L. Hsu, "On the maximum empty rectangle problem," *Discrete Applied Mathematics*, vol. 8, no. 3, pp. 267–277, 1984.
- [14] A. I. Rughani, T. S. M. Dumont, Z. Lu et al., "Use of an artificial neural network to predict head injury outcome," *Journal of Neurosurgery*, vol. 113, no. 3, pp. 585–590, 2010.
- [15] S.-Y. Ji, R. Smith, T. Huynh, and K. Najarian, "A comparative analysis of multi-level computer-assisted decision making systems for traumatic injuries," *BMC Medical Informatics and Decision Making*, vol. 9, no. 1, article 2, 2009.
- [16] B. C. Pang, V. Kuralmani, R. Joshi et al., "Hybrid outcome prediction model for severe traumatic brain injury," *Journal of Neurotrauma*, vol. 24, no. 1, pp. 136–146, 2007.
- [17] M. L. Rohling, M. E. Faust, B. Beverly, and G. Demakis, "Effectiveness of cognitive rehabilitation following acquired brain injury: a meta-analytic re-examination of cicerone et al.'s (2000, 2005) systematic reviews," *Neuropsychology*, vol. 23, no. 1, pp. 20–39, 2009.
- [18] J. Whyte and T. Hart, "It's more than a black box; it's a Russian doll: defining rehabilitation treatments," *American Journal of Physical Medicine & Rehabilitation*, vol. 82, no. 8, pp. 639–652, 2003.
- [19] K. D. Cicerone, D. M. Langenbahn, C. Braden et al., "Evidence-based cognitive rehabilitation: updated review of the literature from 2003 through 2008," *Archives of Physical Medicine and Rehabilitation*, vol. 92, no. 4, pp. 519–530, 2011.
- [20] K. Gibert and A. García-Rudolph, "Desarrollo de herramientas para evaluar el resultado de las tecnologías aplicadas al proceso rehabilitador Estudio a partir de dos modelos concretos: Lesión Medular y Daño Cerebral Adquirido," in Posibilidades de Aplicación de Minería de Datos para el Descubrimiento de Conocimiento a Partir de la Práctica Clínica, Informes de Evaluación de Tecnologias Sanitarias, AATRM núm. 2006/11,

- Cap 6, Plan Nacional para el Sistema Nacional de Salud del Ministerio de Sanidad y Consumo, Madrid, Spain; Agència d'Avaluació de Tecnologia I Recerca Mèdiques, Barcelona, Spain, 2007.
- [21] J. Serra, J. L. Arcos, A. Garcia-Rudolph, A. García-Molina, T. Roig, and J. M. Tormos, "Cognitive prognosis of acquired brain injury patients using machine learning techniques," in *Proceedings of the International Conference on Advanced Cognitive Technologies and Applications (COGNITIVE '13)*, pp. 108–113, IARIA, Valencia, Spain, 2013.
- [22] A. Marcano-Cedeño, P. Chausa, A. García, C. Cáceres, J. M. Tormos, and E. J. Gómez, "Data mining applied to the cognitive rehabilitation of patients with acquired brain injury," *Expert Systems with Applications*, vol. 40, no. 4, pp. 1054–1060, 2013.
- [23] V. Jagaroo, *Neuroinformatics for Neuropsychologists*, Springer, 1st edition, 2009.
- [24] A. Dumitrescu and M. Jiang, "On the largest empty axis-parallel box amidst *n* points," *Algorithmica*, vol. 66, no. 2, pp. 225–248, 2013.
- [25] B. Chazelle, R. L. Drysdale, and D. T. Lee, "Computing the largest empty rectangle," SIAM Journal on Computing, vol. 15, no. 1, pp. 300–315, 1986.
- [26] A. Aggarwal and S. Suri, "Fast algorithms for computing the largest empty rectangle," in *Proceedings of the 3rd Annual Sym*posium on Computational Geometry, pp. 278–290, Waterloo, Canada, June 1987.
- [27] M. McKenna, J. O'Rourke, and S. Suri, "Finding the largest rectangle in an orthogonal polygon," in *Proceedings of the 23rd Annual Allerton Conference on Communication, Control and Computing*, pp. 486–495, Urbana Champaign, Ill, USA, October 1985.
- [28] H. S. Baird, S. E. Jones, and S. J. Fortune, "Image segmentation by shape-directed covers," in *Proceedings of the 10th International Conference on Pattern Recognition*, vol. 1, pp. 820–825, June 1990.
- [29] J. Edmonds, J. Gryz, D. Liang, and R. J. Miller, "Mining for empty spaces in large data sets," *Theoretical Computer Science*, vol. 296, no. 3, pp. 435–452, 2003.
- [30] J. Augustine, S. Das, A. Maheshwari, S. C. Nandy, S. Roy, and S. Sarvattomananda, "Recognizing the largest empty circle and axis-parallel rectangle in a desired location," CoRR, abs/1004.0558v2, 2010.
- [31] J. M. Tormos, A. Garcia-Molina, A. Garcia Rudolph, and T. Roig, "Information and communications technology in learning development and rehabilitation," *International Journal of Integrated Care*, vol. 9, 2009.
- [32] C. S. Green and D. Bavelier, "Action-video-game experience alters the spatial resolution of vision," *Psychological Science*, vol. 18, no. 1, pp. 88–94, 2007.
- [33] M. D. Lezak, Neuropsychological Assessment, Oxford University Press, New York, NY, USA, 3rd edition, 1995.
- [34] L. J. Trettin, Executive functions following traumatic brain injury: the impact of depression upon performance [Ph.D. thesis], Pacific Graduate School of Psychology, 2007.